

小中学校教員の抱える問題解決を目的とした 統計リテラシー教育の提案 —— 仮説検定と結果のまとめ方 ——

音 山 若 穂・山 口 陽 弘
群馬大学大学院教育学研究科教職リーダー講座
(2013年9月18日受理)

A Proposal on Statistical Literacy Education to Help Primary and Lower Secondary School Teachers Work out the Day-to-day Problems : The Ways of Hypothesis Testing and Summarizing the Results

Wakaho OTOYAMA, Akihiro YAMAGUCHI
Program for Leadership in Education, Graduate School of Education, Gunma University
(Accepted on September 18th, 2013)

問 題

今日、心理や教育の研究や実践を進める過程において、統計的な処理は頻繁に登場する。例えば児童・生徒の実態を把握するためにアンケートを行い、人数(度数)や比率(%)を集計したり、テストを行って平均値を求めたりすることは教育現場で日常的に行われている。そして、そこで得た結果から一定の傾向を読み取ったり、どの教授法が効果的かといった検証を行うことも多い。しかし、そうした処理の過程は統計学的な理論に支えられているのであり、知識が無ければ、ときに誤った結果を導くこともありうる。山口(2013)は、教育に関する研究には、概して初歩的な、統計学的知識の欠落からくるミスが多く、特に信頼性と妥当性の2つの概念は、その重要性がしばしば強調されるものの、その割には教育評価の文脈の中で解説されることが少ないと指摘している。

信頼性の高い、妥当な結論を導くためには、研究法の理解と、統計的分析に関する理解の双方が必要である。前者は前稿(山口ら、印刷中)に譲り、本稿では統計的分析に焦点を当てる。

1. 本学教職大学院における 研究報告書の概観

本学教職大学院の発足以来、これまでに提出された課題研究報告書を概観した。

対象は平成21～25年度の課題研究報告書59編(学校運営コース29編、児童生徒支援コース30編)である。このなかで、統計的内容(児童生徒や保護者、教職員などを対象としたアンケートや学力テストなど、何らかのデータを自らもしくは共同で収集し、それを分析または再分析しているもの。既存の統計資料の引用は除く)を含む報告書の数とその概要を表1に示す。

表 1 群馬大学教職大学院における課題研究のうち、統計処理を含む報告の概要

No. ¹⁾	研究テーマ ²⁾	調査の内容、測定 ³⁾	対象（対象者数 n）	分析の概要 ⁴⁾	検定 ⁵⁾	データ ⁷⁾
1	授業改善をめざした組織的な校内研修の取り組み	①校内研修やキャリアキュラママネジメントに関する実態 ②授業研究会で協議した授業参観の観点 60 項目	① 4 市町の小中学校の研修主任 (n=41)、一般教諭 (n=117) ② 勤務校全職員	① 度数、平均値による比較、ニース度（重要度と実現度をもとに算出）の比較 ② 度数の比較		
2	学校における保護者への対応の在り方	①教育現場における保護者との連携体制の構築に関する調査 (a) ②勤務校における部活動に関する調査 ③部活動ガイドブックに関する調査	①県内公立教職員 (n=1874) ②保護者 (n=207)、生徒 (n=225)、教師 (n=17)、 ③児童 (n=64)、保護者 (n=62)	① 度数の比較、校種・周辺環境・性別・年齢・役職と項目とのクロス集計 ② 度数の比較、学年と項目とのクロス集計 ③ 度数の比較、入学予定者・在学者と項目とのクロス集計 ④ 度数の比較、校種・周辺環境・性別・年齢・役職と項目とのクロス集計		
3	学校と地域社会、関係機関との連携体制	保護者との連携体制の構築 (a と同じ)	県内公立教職員 (n=1874)	度数の比較、性別・年齢・役職と項目とのクロス集計		
4	学校危機へ対応するための教職員の協働体制	保護者との連携体制の構築 (a と同じ)	県内 4 都市の小中学校教務主任 (n=86)	度数の比較		
5	教育課程の改善における効果的な組織連携	自校の指導計画の評価改善の仕組みや流れ	県内公立校教職員 (n=1874)	度数の比較、校種・周辺環境・役職と項目とのクロス集計		
6	家庭・地域との連携を重視した学校運営	保護者との連携体制の構築 (a と同じ)	② 教育実践公開サイト運営者 (n=30)	度数の比較、校種・周辺環境・役職と項目とのクロス集計		
7	教育実践セルフアークアイビングの実践と提案	① サイトの検索結果、コンテンツ数、訪問者数等 ② 運営者アンケート ③ 自校における意識調査	② 教育実践公開サイト運営者 (n=30)	度数の比較、校種・周辺環境・役職と項目とのクロス集計		
8	規範意識の育成を目指した指導体制	自校における意識調査	1~3 学年生徒 (n=291)	度数の比較、男女別クロス集計		
9	家庭・地域の学校への参画意識を高めた実践	学校評価アンケート	教職員 (n=21)、保護者 (n=238)	度数の比較、自由記述の一覧		
10	学校現場における教師の保護者への対応力向上	① 自校における意識調査 ② 保護者との連携体制の構築 (a と同じ) ③ 中学校入学期の保護者へのアンケート	① 勤務校教職員 ② 県内公立教職員 (n=1874) ③ 6 年生保護者 (n=90)	① 度数の比較 ② 校種と項目とのクロス集計 ③ 度数の比較		
11	学級の共感性を高める実践研究	多文化的共感性尺度 (登録, 2003)	クラスの生徒 (n=33)	学級と、登録 (2003) の結果との平均値比較、4 月と 6 月、9 月と 11 月の 2 サイクル、および 4 時点の推移の検討	○	
12	学校における同僚性・協同性の形成に関する一考察	佐藤・児島 (2007) を元で作成したアンケート	教職員 (n=20)	度数の比較、3 月と 12 月の 2 時点間の比較グラフ		
13	中 1 ギャップを解消するための小中学校の連携	① 中 1 ギャップ要因に関するアンケート ② 中 1 ギャップの実態アンケート	① 小 6 児童 (2 校: n=85、n=24) ② 中 1 生徒 (実習校 n=63、勤務校 n=24)	① 度数の比較、2 つの小中学校をそれぞれ分析 ② 現在の状況と、入学期を思い出し評価したものとの度数の比較 (グラフ)、男女別比較 (グラフ) ③ 重要度、実現度、ニース度 (重要度と実現度をもとに算出) の比較 ④ 度数の比較		
14	児童の基本的生活習慣の確立と規範意識の向上を目指して	① 学校教育に関わる重要度・実現度調査 ② 校内研修実施後のアンケート ③ C&S 質問紙 ④ 学校評価アンケート等	①、②ともに教職員 (n=40) ③ 5 年生 4 クラスの児童 ④ 教職員、保護者	① 自由記述の列記、および参考になった割合 (%) ② クラスごとの散布図 ③ 度数の比較 ④ 現在の状況と、入学期を思い出し評価させたものとの度数の比較 (グラフ) ⑤ 重要度、実現度、ニース度 (重要度と実現度をもとに算出) の比較		
15	基礎学力の定着を目指した指導体制の構築	① 中 1 ギャップを握ること ② 校内研修への取り組みに関するアンケート	① 中 1 生徒 (22・23 年度生) ② 学年教職員 (n=6)	① 現在の状況と、入学期を思い出し評価させたものとの度数の比較 (グラフ) ② 度数の報告 ③ 度数の報告「授業について会話を行う頻度」「専門教科が異なる教員と授業研究会以外の場面で語る頻度」について、5 月と 12 月との間で入数の増減比較		
16	中学校において専門教科の指導力を高めるための学年会の在り方	課題研究に関するアンケート (5 月と 12 月の 2 回実施)	学年教職員 (n=7)	① 道徳授業担当者 (n=19) ② 全児童 (1~6 年生) ③ 1 クラスの児童 (n=30) ④ 教職員 (n=28)	①、③、④ 度数の比較。 ② グラフのみ (度数の表示はなし)	
17	効率的な校内研修の進め方	① 道徳についてのアンケート ② 道徳の内容項目についての自己評価 ③ 授業の事前アンケート ④ 校内研修に関するアンケート	① 道徳授業担当者 (n=19) ② 全児童 (1~6 年生) ③ 1 クラスの児童 (n=30) ④ 教職員 (n=28)	①、③、④ 度数の比較。 ② グラフのみ (度数の表示はなし)		
18	授業力を高める校内研修のあり方に関する研究	① ワールドカフェの有効性、自分の授業力の良きや課題、目標に確立した評価の効果等 (8 月調査) ② 多忙感について等 (11 月調査)	いずれも教職員 (n=18)	①、②とも自由記述中心だが、一部、度数の報告もみられる		
19	専門学校における地域連携モデルの構築	① 学校生活の満足度などのアンケート ② 地域連携に関する生徒アンケート	① 平成 20 年度卒業生 ② 商業科生徒 (n=118)	① 度数の比較、工業系と商業系の比較 ② 度数の比較、調査研究グループと、その他の 2 群間の比較 (グラフ)		
20	自己有用感を高め、学校生活の充実への意欲を育む学年学級経営の在り方	Q-U テストを 6 月および 11 月の 2 回実施	1 クラスの生徒 (n=31)	① 度数の比較、工業系と商業系の比較 ② 度数の比較、調査研究グループと、その他の 2 群間の比較 (グラフ)		
21	学校・家庭・地域の連携をもとに進めるキャリア教育	キャリア教育に関するアンケート	教職員 (n=23)	① 度数の比較、工業系と商業系の比較 ② 度数の比較、調査研究グループと、その他の 2 群間の比較 (グラフ)		
22	教師の授業力向上を図る学校体制と同僚性の構築	授業力向上に関するアンケート (重要度・達成度)	教職員 (n=23)	① 度数の比較、工業系と商業系の比較 ② 度数の比較、調査研究グループと、その他の 2 群間の比較 (グラフ)		
23	通常学級における LD の疑いのある児童の読み書きの支援と校内特別支援教育体制づくり	市販テストの得点の、抽出現と平均点との差	市販テストの生徒 (n=31)	① 度数の比較、工業系と商業系の比較 ② 度数の比較、調査研究グループと、その他の 2 群間の比較 (グラフ)		
24	小学校国語科における話し合いを深めるための学習指導	授業の事前、事後におけるアンケート	3 クラスの児童 (事前 n=89、事後 n=88)	① 度数の比較、学級別平均回客数の実施時期 (3 時点) の推移、自由記述のカテゴリ一化についての一致率 (2 時点間比較)、カテゴリ一化した記述数のカテゴリ一間比較、各学級のテスト平均点と、全体的な得点分布のグラフ		

25	学級経営を充実させる学級活動の在り方	①振り返りカード ②算数50問テスト、算数まとめテスト ③学級の雰囲気質問紙	1 クラスの児童 (n=34)	①度数の比較 ②学級平均と全国平均の比較 ③下位2位ごとの学級の1値(4月と12月の2時点)		
26	学校における食育に関する実践研究	中村ら(2007)による5段階到達度評価	1 クラスの児童 (n=27)	度数の比較		
27	外国籍児童の在籍する学級における望ましい算数科指導	①算数の学習に関するアンケート調査 ②文章題に関する事前・事後テスト ③文章題に関する事後テスト	いずれも1クラスの児童 (n=25)	①度数の比較 ②事前・事後それぞれ、正答者数と誤答の種類別の人数		
28	生徒一人一人の自己肯定感を支える授業づくり	自己肯定感尺度(平石,1990)の対人領域7項目、およびアンケート	生徒 (n=91)	自己肯定感尺度、およびアンケート項目について、4月と10月の2時点間の比較(対応のあるt検定)。尺度のα係数。	○	○
29	中学校理科における科学的思考力を高める指導法	国立教育政策研究所(2005)を元にしたアンケート調査(事前・事後)	2 クラスの生徒 (n=56)	事前・事後における度数の比較(カイ2乗検定・残差分析)	○	
30	望ましい人間関係を育てる学級活動	振り返りのアンケート	1 クラスの生徒 (n=32)	度数の比較		
31	学習内容の系統性を踏まえた小学校理科の学習指導	①授業の事前・事後におけるアンケート。 国立教育政策研究所(2007)を元にして作成。 ②市販の单元テスト	いずれも第5学年 (n=128)および第6学年 (n=150)	①事前と事後の度数比較(Friedman検定)をして検定統計量等(不明)。第5学年は事後は1回測定、第6学年では事後は2回(事前1回)測定。 ②クラス別、時点の平均値の比較	○	
32	わかる授業により児童の学習意欲を高める社会科学習指導	①児童の振り返りの自由記述 ②縦割り班活動の実施状況	①、②ともに1クラスの児童 (n=27)	①度数の比較、事前と事後の比較グラフ ②記述内容を分類し、度数の比較		
33	高学年児童のリーダーシップを育成する指導の工夫	①引継情報資料(情動、認知等の発達状況) ②ビデオ分析によるエピソード形態(藤崎ら1985)	県内全小学校 (n=205)	度数の比較。結果の表には毛利(2002)による全国調査の結果を併記。		
34	協同的に遊ぶ姿を支える保育	CRT-II学力検査の結果	①5歳児 (n=33) ②抽出児	①抽出児の平均と全体平均との比較 ②4、6、10月の3時点での該当数の比較(クロス表)		○
35	英語でコミュニケーションが進んでできる生徒の育成	①事前アンケート ②授業ワークシート ③事後アンケート	1 年生 (n=30)	得点率		
36	中学生の規範意識を育てる道徳の時間	①授業の事前アンケート ②授業ワークシート ③事後アンケート	いずれも1クラスの生徒 (n=40)	①度数の比較。場面ごとの度数比較(グラフ) ②事前アンケートと、授業ワークシートとの間の度数の比較(グラフ)		
37	小学校国語科における「読むこと」に生かす文法指導	「つなぎ言葉」に関する新井(2004)のアンケート	1 クラスの児童 (n=29)	③場面ごとに、事前・事後間の度数比較。McNemar検定の有意確率のみ表示		
38	小学校体育科におけるボール運動技能の向上	授業の事前アンケート	1 クラスの児童 (n=33)	設問ごとの正答率を事前・事後の2時点で比較。 度数の比較。		
39	式、計算の意味理解を深める算数科指導法	①事前アンケート ②市販の单元テスト ③国立教育政策研究所(2006)の「計算に関する力」式や計算の意味を理解することの5題 ④多田ら(2009)を元にして作成した16項目のアンケート	いずれも1クラスの児童 (n=39)	①度数の比較 ②実践群、担任担当群、学級全体との3群で平均点の比較。 ③設問ごとに、上記3群および全国間で正答率を比較。6月と11月の正答率推移グラフ(実践クラス)。 ④質問項目ごとに分析。いずれも実践クラスと担任クラスとのそれぞれにおいて、事前事後の2時点間の比較をしている(対応のあるt検定)	○	○
40	論理的に思考するための言葉を育む中学校国語科学習指導	①全国学力・学習状況調査を基にした検証テスト ②古典および授業内容の好感度アンケート	4 クラスの生徒 (n=133)	①設問ごとに全国と対象校それぞれの反応率、およびその差 ②度数(%),および事前と事後の%差		
41	算数教育	LDI-R	抽出生徒 (n=3)	粗点合計、グラフを含むプロフィール表を掲載		
42	児童が実感し、断に落ちることを目標とする小学校社会科	群馬県教育委員会(2011)「ぐんまの子ども」の基礎・基本習得状況調査」を元にして作成した5項目	1 クラスの児童 (n=20)	度数の比較。5月と11月の結果をそれぞれ示している。その差についての数値的分析結果は特に示されていない。		
43	生徒の自治的能力を高める特別活動の取組	①日常の係活動に関する自己評価8項目 ②文化祭の係活動に関する自己評価8項目 ③NRTを参考に作成した「読むこと」の力 ④犬塚(2002)などを参考に作成した「読むこと」の17項目 ⑤NRTを参考に作成した「読むこと」の力 ⑥事後の保護者アンケート(自由記述) ⑦事後の児童自己評価(3項目)	いずれも1クラスの生徒 (n=21)	①項目ごとにクラス平均値を求めグラフ表示(数値なし)。1学期と2学期それぞれ、中間と期末の平均値を比較(グラフ)。 ②日常の係活動と文化祭の係活動の差の比較(対応のあるt検定)	○	○
44	国語科における読解力の育成	①日常の係活動に関する自己評価8項目 ②文化祭の係活動に関する自己評価8項目 ③NRTを参考に作成した「読むこと」の力 ④犬塚(2002)などを参考に作成した「読むこと」の17項目 ⑤NRTを参考に作成した「読むこと」の力 ⑥事後の保護者アンケート(自由記述) ⑦事後の児童自己評価(3項目)	いずれも1クラスの生徒 (n=27)	①項目ごとに、4月～12月の5時点における平均値の推移を検討(グラフ)。分散分析を行っている(統計量の表記なし) ②対象校と全国との正答率比較。	○	
45	支え合う学級集団を目指して	①小学生用SEL-8S(小泉ら,2011) ②「できてほしいこと」自己評価(13項目)	1 クラスの児童 (n=28)の児童および保護者	①主な記述の列記 ②度数の比較		
46	児童の社会性を育む指導の工夫	①安永(2006)を参考に作成した33項目 ②「できてほしいこと」自己評価(13項目)	いずれも1クラスの児童 (n=30)	①下位2位ごと、3時点の平均値の比較(分散分析および多重比較) ②項目ごと、2時点の平均値の比較(対応のあるt検定)	○	○
47	伝え合う力を育てる中学校国語の学習指導	①安永(2006)の取り組みアンケート	いずれも1学年3クラスの生徒 (n=86)	①尺度ごと、4時点の平均値の比較(グラフ)。分散分析(統計量の表記なし) ②度数の比較	○	
48	食領域の基礎的知識や技能を習得させる家庭科指導法	「ぐんまの子ども」の基礎・基本習得状況調査」を元にして作成した①アンケート14項目、②テスト問題9題	いずれも4クラスの生徒 (n=137)	①度数の比較(グラフ)。および4月と12月の差の比較。対応のあるt検定で有意差がみられた項目が示されている(統計量の表記なし) ②当該校と全県との間で正答率の比較	○	

*1 22までの運営23以降が支援コース
*2 報告事項欄目のうち、一部語句を省略してある
*3 すでに公開されている他の調査結果、データ等を単に引用したものは除外した
*4 度数の比較には、比率(%)によるものも含む
*5 統計的仮説検定の結果が示されているもの
*6 平均値が示されている結果のうち、標準偏差(SD)が示されているもの
*7 平均値が示されている結果のうち、粗データが表示されているもの

統計的内容を含む報告書の数、59編中48編(81.4%)、学校運営コース22編(75.9%)、児童生徒支援コース26編(86.7%)であり、およそ8割の報告書が何らかの統計的内容を含んでいた。

調査(測定)の対象は児童生徒、教職員、保護者それぞれであるが、大別して1回の調査による実態把握と、繰り返し調査を行って継続的变化を検討するものとがみられる。対象者数(n)はばらつきが大きいものの、多くが10人以上のデータを確保している。

統計的内容では、事例検討の1編を除きほぼ全ての報告書が人数など度数(%表示など、比率を含む)を扱っており、平均値を扱うものは12編であった。平均とともに散布度も示しているものは5編、粗データが付属していて再分析可能なものは2編であった。

統計的検定を用いたものは9編あり、対応のある t 検定が4件、繰り返しのある分散分析が3件、 χ^2 検定が2件であった。ただし、検定統計量が見当たらないなど情報が不足しているものも含まれた。多変量解析を扱っているものはなかった。

以上のように、多くの報告が何らかのデータを収集し、ある程度のデータ数を確保していて推測統計的な処理も可能であるにも関わらず、多くが度数の記述にとどまっているか、統計的処理や結果の表示が不足していて物足りないものが少なくない。

この理由の一つとして、仮説検定に関わる統計リテラシー、なかでも具体的な分析手順や結果の表記の仕方など実証的な知識が不足していることが考えられる。例えば本学教職大学院で用いられる主要な参考書の一つである佐藤ら(2013)をみると、教育評価における統計学的知識の必要性についての理論的な解説にかなりの紙幅が割かれている。読者は、ここから実践研究において必要な統計処理の基本的な考え方や注意点を知ることができるが、検定の仕組みや手順、結果の書き方についてはあまり触れられていない。

しかし、少なくとも本学教職大学院において多くの報告が統計的内容を含み、そのおよそ2割は統計的検定を用いている現状を踏まえれば、統計処理に

関するある程度具体的な解説を行うことは必要であろう。その際に重要なのは、現職教員が負担なく活用できるよう、実践的内容にポイントを絞ることであろう。

そこで本稿では、これまでの報告書で多く扱われてきた事前・事後測定のケースを念頭に、対応のある t 検定、1要因被験者内分散分析について、解説例を示すこととした。

2. 事前・事後の2時点の比較

実践をはさみ事前と事後とで、同じ項目のアンケートを行い、両者の差から実践の効果を示したい場合を考える。以下の例は、小3の1クラス30名を対象に学級活動でコミュニケーションスキル訓練などの実践を行った結果の一部を改変したものである(高橋, 2013)。項目は、担任が学級目標などを踏まえて独自に作成した「身につけて欲しい」ことである。オリジナルは13項目あるが、今回はうち3項目と、13項目の合計値のみ示す(付表1)。

(1) Excel の分析ツールでの分析

Excelで、1行目には変数名、2行目以下に1行あたり1名のデータを入力する。対応のある t 検定では、事前と事後のデータが揃っていることが前提なので、欠席などどちらかが欠落している場合は事前に除外する必要がある(今回は1名分を除外)。

	A	B	C	D	E	F	G	H	I
1	No	事前合計	事前1	事前2	事前3	事後合計	事後1	事後2	事後3
2	1	27	3	4	3	45	3	4	4
3	2	36	4	4	4	47	4	4	4
4	3	40	4	4	4	42	3	4	3
5	4	37	3	4	2	41	4	4	2
6	5	42	3	4	2	51	4	4	4
7	6	37	3	4	1	45	3	3	4
(中略)									
27	26	33	3	4	2	49	4	4	4
28	27	27	3	3	2	38	3	2	3
29	28	44	2	4	4	43	4	4	3
30	29	26	2	3	1	45	4	4	4
31	30	39	3	4	2	39	3	2	2

図2 Excelでのデータ入力例

データ入力後、[ファイル]→[オプション]→[アドイン]→[設定]で[有効なアドイン]の欄にある[分析ツール]のチェックを入れて[OK]をクリック

クする。この操作の後で「データ」タブをクリックすると、右端に「データ分析」が表示されるのでクリックする。今回使用するのは、分析ツールの中の「t 検定：一対の標本による平均の検定」である。

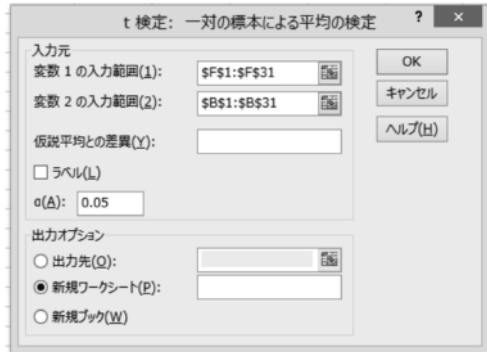


図3 分析ツールの設定画面の入力例

分析は、項目ごとに行う。変数1の入力範囲に「\$F\$1:\$F\$31」、変数2の入力範囲に「\$B\$1:\$B\$31」と入力またはマウスで範囲指定する。入力範囲の第1行目は「事前合計」「事後合計」のように変数名(ラベル)なので、「ラベル」のチェックを入れる。他の欄は設定しなくてよい。以上で「OK」をクリックする(図3)。

結果を図4に示す。事前・事後、それぞれ13項目の合計点の平均値が4行目に示されている。

	A	B	C
1	t 検定：一対の標本による平均の検定ツール		
2			
3		事後合計	事前合計
4	平均	44.6	38.2
5	分散	34.66206897	48.44137931
6	観測数	30	30
7	ピアソン相関	0.44129519	
8	仮説平均との差異	0	
9	自由度	29	
10	t	5.116566583	
11	P(T<=t) 片側	9.17278E-06	
12	t 境界値 片側	1.699127027	
13	P(T<=t) 両側	1.83456E-05	
14	t 境界値 両側	2.045229642	

図4 分析ツールの分析結果

5行目「分散」(S^2)は、ちらばりの大きさを示す指標で、一人ひとり($i=1, 2, 3, \dots, n$)の値(X_i)が平均値(\bar{X})の近くに集まっていれば小さく、 X_i がばらばらで、 \bar{X} から離れた値が多ければ大きくなる。 N をデータ数とすると、次式により求められる。

る。

$$S^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1} \quad (2.1)$$

分散の平方根を取ったものが標準偏差(S , Standard Deviation; SDとも略される)である。報告書等に結果を記す際には、 \bar{X} だけでなく、 S^2 ないし S も合わせて示すことが一般的である。

$$S = \sqrt{S^2} \quad (2.2)$$

(2) 「仮説平均との差異」と標本抽出

ところで、対応のあるt検定においては、まず一人ひとりについて事後と事前の差 D_i を求める。事前の値を A_i 、事後の値を B_i として、

$$D_i = B_i - A_i \quad (2.3)$$

としたとき、この D_i の平均 \bar{D} について考えてみる。

かりに事前と事後に差がなく、全員について

$$A_i = B_i$$

であるなら、 $\bar{D}=0$ になる。例えば、特に工夫した授業など行わず、同じクラスに単に2度測定を行っただけならば、 $\bar{D}=0$ となることが期待される。逆に言えば、何か工夫した授業をしたとして、その効果を示すためには、まず $\bar{D}=0$ ではないことを確かめる必要がある。このような場合は、「仮説平均との差異」は0(すなわち何も入力しなくてよい)である。

ところで、今回の例では、 $\bar{D}=6.40$ であり、平均してこれだけ事後の値が伸びている。ただし、ここから授業の効果があったと直ちに言えるかという、必ずしもそうではない。実際には特に何も行わなかったとしても、測定誤差もあったりして、全員について $A_i=B_i$ になるとは限らず、 $\bar{D}=0$ とならない場合もありうるからである。そこで、 $\bar{D}=0$ であるとする仮説の真偽については、もう少し考えを進める必要がある。

いま、ある大きな集団(例えば学校全体)のデータがあって、平均 μ_x 、標準偏差 σ_x であるとする。こ

の集団を母集団と呼ぶ。ここから、抽選で（無作為に） n 人分のデータを取り出すことを考える。選ばれた n 人を標本と呼ぶ。そこでこの n 人の平均、すなわち標本平均 \bar{X} を求める。

以上の作業を何度も繰り返すと、繰り返し数だけ標本平均 \bar{X} が得られる。標本に誰が選ばれるかは毎回異なるだろうから、 \bar{X} にもちらばりが生じる。このようにして多くの \bar{X} を得たとき、その分布（ \bar{X} の標本抽出分布と呼ばれる）にはどのような特徴があるのかを考えてみる。

\bar{X} の分布の平均を $\mu_{\bar{X}}$ 、標準偏差を $\sigma_{\bar{X}}$ とすると、母集団との間に以下の関係があることが知られている。

$$\mu_{\bar{X}} = \mu_X \quad (2.4)$$

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} \quad (2.5)$$

さらに、母集団が正規分布であれば \bar{X} もまた正規分布であること、標本の大きさ（ n ）が大きければ、（母集団の分布によらず）正規分布に近づいていくことが知られている（山内 2009；第 7 章が分かりやすい）。

(3) 標準正規分布と確率

正規分布とは図 5 のような分布であり、次式で定義される。

$$Y = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (2.6)$$

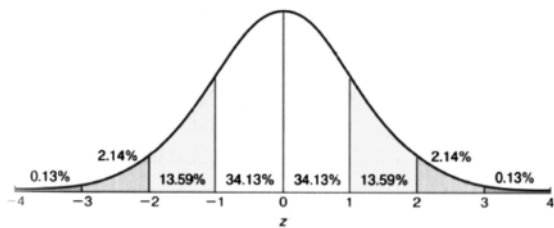


図 5 正規分布と z

分布（山の形）は μ と σ で決まるが、 $\mu=0$ 、 $\sigma=1$ とすると、分布はひとつに決まる。これを標準正規

分布と呼ぶ。

いま、ある集団テストを行ったとして、一人ひとりの得点 X_i を次式により z_i に変換する。

$$z_i = \frac{X_i - \bar{X}}{S} \quad (2.7)$$

z_i の平均は 0、標準偏差は 1 である。 z_i は標準得点とも呼ばれる。なお、これを以下の式により解釈しやすくしたものが「偏差値」である。

$$\text{偏差値} = 10z + 50 \quad (2.8)$$

統計法のテキストには、巻末に標準正規分布表が掲載されており、分布の全体を 1 としたとき、 z_i 以上もしくは z_i 以下にそれぞれどれだけの割合が含まれるのかを求めることができる。

たとえば受験者 10 万人、平均 50、標準偏差 10 の全国テストがあったとする。このテストで 70 点であった生徒 i については、

$$z_i = \frac{70 - 50}{10} = 2.0 \quad (2.9)$$

である。成績の分布が正規分布するとすれば、標準正規分布表から、この生徒より上位には $p=0.0228$ (2.28%)、下位には $p=0.9772$ (97.22%) が含まれることが分かる。すなわちこの生徒より上位の生徒は約 2 千人いる計算となる。

(4) 帰無仮説と仮説の検定

元の例に戻ろう。いま、母集団 A から無作為に標本抽出を繰り返し、 \bar{X} の分布を得たとして、 $\mu_{\bar{X}}=0.0$ 、 $\sigma_{\bar{X}}=2.0$ であったとする。そこに、1 つの標本データ（標本 B）があり、その平均は 4.0 であったとする。標本 B が母集団 A から抽出された標本であると言えるかどうかを、統計的検定によって判断したい（検定の考え方については多くの平易な解説書がある。例えば服部ら 1996、古谷野ら 1992 など）。

この場合、まず「標本 B は、母集団 A から無作為に抽出された標本のひとつである」という仮説を立てる。この仮説を帰無仮説 (H_0) と言う。この逆、「標本 B は、母集団 A から無作為に抽出された標本

のひとつではない」は対立仮説 (H_1) と呼ばれる。

標本抽出分布は正規分布とみなせるはずであるから、かりに H_0 が正しいなら、標本 B の平均を次式により変換し

$$z = \frac{\bar{X} - \mu_x}{\sigma_{\bar{X}}} = \frac{4.0 - 0.0}{2.0} = 2.0 \quad (2.10)$$

標準正規分布表から、4.0 以上の平均をもつ標本が出現する確率を求めることができる。結果は $p = .0228$ であり、「標本 B は、母集団 A から無作為に抽出された標本のひとつである」(つまり、 H_0 が真) と考えるには、この確率はきわめて小さい。

統計的検定では、あらかじめこの確率に一定の水準を設け、その水準以下であれば、 H_0 を棄却する。この水準を有意水準 (α) と呼び、一般に $\alpha = 0.05$ または $\alpha = 0.01$ のように低い値に設定する。

\bar{X} のばらつきは、 μ_x からみて正の方向と負の方向の双方が考えられるので、実際には α を 1/2 ずつ正負双方に振り分ける。これを両側検定と呼ぶ。 α を振り分けない片側検定という手法もあるが、両側検定のほうが一般に考えやすい。

元に戻って、今回の例題について考えてみる。まず、帰無仮説は「事前と事後とに差はない ($\bar{D} = 0$)」とする。すなわち、 \bar{D} は $\mu_x = 0$ である母集団から無作為抽出された標本平均であると考ええる。

そこで検定を行いたい、そもそも標本は 1 つ (1 クラス分の結果) しかなく、何度も調査を行ったわけではないので、 $\sigma_{\bar{x}}$ が分からない。母集団 (たとえば小 3 児童全体) の分散 σ_x が分かれば、(3.2) 式から $\sigma_{\bar{x}}$ を求めることができるが、これも分からないため、結局 z が計算できず、上述の手順が使えない。

このような場合、 z に代わって、ステューデントの t という統計量を用いる方法が知られている。 t は、 σ_x の代わりに、 s を用いる。 $\mu_x = \mu_x = 0$ であるから、

$$t = \frac{\bar{D} - \mu_x}{S_{\bar{D}}} = \frac{\bar{D} - 0}{S_{\bar{D}}} \quad (2.11)$$

ただし、

$$S_{\bar{D}} = \frac{S_D}{\sqrt{n}} \quad (2.12)$$

S_D は D の標準偏差である。

t を求めると、図 4 の 10 行目の値に相当する。

$$t = \frac{6.400 - 0}{6.851 / \sqrt{30}} = 5.117 \quad (2.13)$$

t も同様に、統計法のテキストの巻末に表が掲載されている。ただし t から確率を求めるには、次式によって求められる自由度 df の値が必要である。

$$df = n - 1 = 30 - 1 = 29 \quad (2.14)$$

n は D のデータ数、すなわち事前・事後のデータが揃っているものの数 (今回は 30) である。

確率値は図 4 の 13 行目、「 $P(T \leq t)$ 両側」の欄に示されている。この値が α 以下であれば、 H_0 は棄却される。 $\alpha = 0.05$ とすると、今回は $p < 0.05$ であり H_0 棄却である。すなわち、今回の標本は、平均 \bar{D} が 0 から大きく (プラス方向に) 離れていて、 $\mu = 0$ の母集団から無作為抽出された標本とは考え難いというのである。つまり、「事前と事後とに差はない」という仮説は棄却されたことになり、結果的に「何らかの差がある」ことになる。これが「統計的に有意差がある」という状態であり、その理由の一つに授業実践の効果が関与していることが示唆されるのである。

(5) 結果のまとめかた

結果のまとめ方の例を表 2 に示す。表 2 では、全項目合計に続き、下位項目 1 ~ 3 も同様に分析し、一覧表にまとめている。

事前と事後それぞれの平均と SD が示されているが、検定ではこれらの差 (例えば全項目合計では 37.86 と 44.89 の差) ではなく、一人ひとりについて D を求めて分析している点に注意が必要である。したがって、表 2 には、 \bar{D} と (2.12) 式で求められる標準誤差 $S_{\bar{D}}$ (Standard Error ; SE) を示した。 \bar{D} を $S_{\bar{D}}$ で除したものが t であるから、ここでは SD の代わりに表示した。 t とともに、 df と p を合わせて表示す

表2 「身に着けてほしいこと」の事前・事後比較

質問項目	事前(A)		事前(B)		差 D (D=B-A)	SE	t	df	p
	平均	SD	平均	SD					
全項目合計	38.20	6.96	44.60	5.89	6.40	1.251	5.117	29	0.000
1 聞いている人を見て、話すことができましたか	3.13	0.82	3.50	0.68	0.37	0.189	1.943	29	0.062
2 相手にきこえる声で話すことができましたか	3.70	0.60	3.57	0.82	-0.13	0.150	-0.891	29	0.380
3 相手に話すときに、笑顔で話すことができましたか	2.80	1.03	3.50	0.82	0.70	0.236	2.971	29	0.006

(以下略)

る。 p は図4の13行目の値だが、指数表示なので「セルの書式設定」で数値を選び、小数点以下の桁数を適宜設定するとよい。 P には、有意差の有無を見分けやすくするために印をつけてある。 $p < 0.05$ の場合に「*」、 $p < 0.01$ の場合に「**」とすることが多いようである。ただし、 p が小さいからといって、効果の大きさが大きいとは限らないので注意が必要である（鋤柄，2002；p.106）。

(6) SPSS による方法

[ファイル] → [開く] → [データ] に進み、[ファイルの種類]を「Excel」に指定することで、excel デー

タを SPSS に読み込める。なお、SPSS では欠損値があっても分析から自動的に除外されるので、入力段階で省く必要は必ずしもない。

読み込み後、[分析] → [平均の比較] → [対応のあるサンプルの t 検定] に進み、[対応のある変数] に分析対象変数を指定し [OK] をクリックする（図6）。

出力結果のうち「対応サンプルの検定」の一番右にある「有意確率（両側）」の値が、 α （0.05 または 0.01）以下であれば「有意差あり」である（図7）。

3. 3 時点以上の比較

1 要因被験者内分散分析を用いる。以下の例は、前述の児童たちを対象に、SEL-8S を 3 時点（ $t_1 \sim t_3$ ）で測定した結果のうち「対人関係」である（付表2）。

(1) Excel の分析ツールでの分析

[データ分析] → [分散分析：繰り返しのない二元配置] を選ぶ。入力範囲に「\$A\$1：\$D\$32」と入

図6 SPSS「対応のあるサンプルの t 検定」設定例

対応サンプルの検定

	対応サンプルの差					<i>t</i> 値	自由度	有意確率 (両側)
	平均値	標準偏差	平均値の 標準誤差	差の 95%信頼区間				
				下限	上限			
ペア 1 事後合計—事前合計	6.400	6.851	1.251	3.842	8.958	5.117	29	.000

図7 SPSS での対応のある t 検定の出力

力またはマウスで範囲指定する。[ラベル]のチェックを入れ [OK] をクリックする (図 8)。

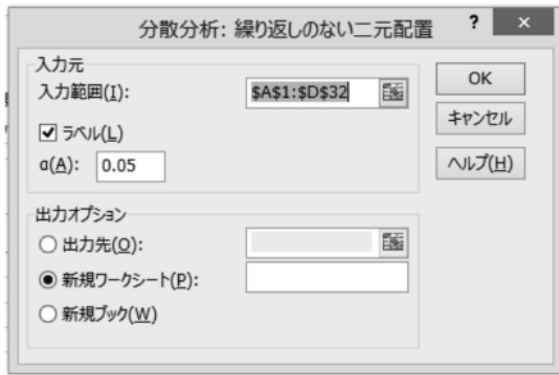


図 8 分析ツールの設定画面の入力例

結果を図 9 に示す。出力のうち重要なのは「分散分析表」である。

(2) 結果の解釈とまとめ方

まず、変動欄から見ていく。47 行目の「合計」の変動 (全平方和とも言う) は、次式で求められる。

$$SS_t = \sum_j \sum_i (X_{ji} - \bar{T})^2 \quad (3.1)$$

\bar{T} は、3 時点を込みにした全てのデータ (31 (人) × 3 (回) = 93 (件) のデータ) の平均で、本例では $\bar{T} = 10.118$ である。なお、 SS_t を自由度 $df = 92$ で除したものが、全データの分散 S^2 であり、(2.1) 式に一致する。

「行」の変動は次式で求められる。

41	分散分析表						
42	変動要因	変動	自由度	分散	観測された分散比	P-値	F 境界値
43	行	193.5323	30	6.451075	7.5895	1.86E-11	1.649141
44	列	20.66667	2	10.33333	12.15686	3.69E-05	3.150411
45	誤差	51	60	0.85			
46							
47	合計	265.1989	92				

図 9 Excel で出力される分散分析表

$$SS_s = \sum_i n_j (\bar{S}_i - \bar{T})^2 \quad (3.2)$$

n_j は時点の数、 \bar{S}_i は 3 時点を込みにした i 番目の被験者 (児童) の平均である。例えば児童 1 についてみると

$$n_i (\bar{S}_i - \bar{T})^2 = 3(6.667 - 10.118) \quad (3.3)$$

となる。同様に全員分を計算し、総和を求めたものが SS_s である。これは個人差が大きいほど、すなわち時点の差を無視したときに、 \bar{T} を基準として \bar{S}_i のちらばりが大きいほど、大となる。

なお、 SS_s の自由度 df_s は、被験者数を n_i として次式で求められる。

$$df_s = n_i - 1 \quad (3.4)$$

今回は 31 人であるので $df_s = 31 - 1 = 30$ となる。

「列」の変動は、

$$SS_A = \sum_j n_i (\bar{A}_j - \bar{T})^2 \quad (3.5)$$

で得られる。 n_i は被験者数、 \bar{A}_j は全被験者を込みにした j 時点の平均である。例えば時点 1 では

$$n_i (\bar{A}_j - \bar{T})^2 = 31(9.484 - 10.118)$$

であり、他の 2 時点についても同様に求めたものを合計すると SS_A が得られる。これは、時点間の差が大きいほど、すなわち個人差を無視したときに、 \bar{T} を基準として \bar{A}_j の差が大きいほど、大となる。

なお、 SS_A の自由度 df_A は、時点の数を n_i として次式で求められる。

$$df_A = n_i - 1 \quad (3.6)$$

今回は 3 時点であるので $df_A = 3 - 1 = 2$ となる

「誤差」(残差とも呼ぶ) の変動はこの場合、3 時点を込みにした各被験者の平均 AS_i のまわりの変動 (個人内変動) を表す。誤差は、個人差としても、時点間の差としても説明されない残りの変動を示すものである。

$$SS_w = \sum_j \sum_i (X_{ji} - \bar{S}_i)^2 \quad (3.7)$$

そして、3 つの変動には以下の関係がある。

$$SS_t = SS_s + SS_A + SS_w \quad (3.8)$$

自由度についても同様の関係がある。

$$df_t = df_s + df_A + df_w \quad (3.9)$$

いま、分析の目的は時点間の差であるから、注目すべきは「列」の変動である。「列」の変動 (SS_A) を自由度で除したものが分散 (または平均平方, Mean Square; MS) であり、「列」の分散 MS_A を「誤差」の分散 MS_w で除したものが「列」における「観測された分散比」 F になる。

$$F = \frac{MS_A}{MS_w} \quad (3.10)$$

MS_A は時点間の差であり、 MS_w は誤差であるから、 H_1 が真であれば、 F の値は大となる。つまり F の値が大きいほど、時点間の効果は偶然生じたものではないという確率が高まると考えることができる。効果が全くないならば、 $F=1$ となる。

そこで、 F の値に一定の基準を設け、その値以上であれば、 H_0 を棄却する。 p は、 H_0 のもとで、この F が得られる確率を示している。この値が α 以下であれば、「有意差あり」と判断することになる。

以上のように、この分析では、被験者の効果を除外した上で、測定時点の効果のみを分析しているという特徴がある。すなわち「行」については F を求める必要はない。そこで、分散分析表は表 3 のように書き直すとよい。

(3) SPSS による方法

[分析] → [一般線形モデル] → [反復測定] を選ぶ (図 10)。「反復測定の因子の定義」では、[水準数] に、測定時点の数 (本例では 3) を入力し [追加] をクリックする (図 11)。

[反復測定] ウィンドウでは、[被験者内変数] に分



図10 SPSS による 1 要因被験者内分散分析



図11 反復測定の因子の定義

析したい変数 (時点 1, 時点 2, 時点 3) を指定し、[オプション] をクリックする (図 12)。

[平均値の表示] 欄に「factor1」を指定し、[主効果の比較] にチェックを入れ、[信頼区間の調整] を選ぶ (この場合は、Bonferroni を選んでいる; 図 13)。
[続行] → [OK] で出力が表示される。

SPSS の出力は、Excel の出力 (図 9) と大きく異なり、いくつかの表から値を拾って、分散分析表を完成させる必要がある。分散分析表 (表 3) に対応する SPSS の出力は、被験者間 (個人差) が図 14 の①、

表 3 「対人関係」の分散分析表

分散分析表					
変動要因	SS	df	MS	F	P
被験者間 (個人差)	193.53	30	6.451	7.590	0.000
被験者内 (時点)	20.67	2	10.333	12.157	0.000
誤 差 (残差)	51.00	60	0.850		
合 計	265.20	92			



図12 反復測定

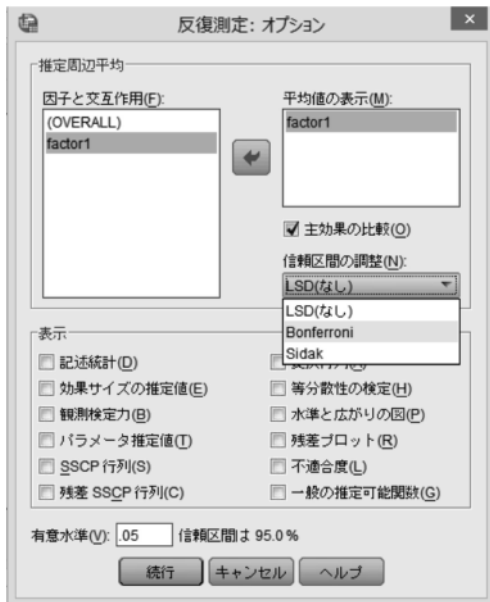


図13 反復測定のオプション

被験者内(測定時点)が図15の②、残差が③である。全体の平方和と自由度は出力されない。(3.8)式と(3.9)式から自分で計算する必要がある。

被験者間効果の検定

測定変数名: MEASURE_1...

	タイプIII平方				
切片誤差	9521.301 193.532	1 30	9521.301 6.451	1475.925	0.000
①	-----	-----	-----		

図14 被験者間効果の検定

(4) 区間推定とグラフ

SPSS では区間推定 (95%信頼区間) の結果も表示されるので (図16), 結果にも含めておくとよい。95%信頼区間は, 次式で表すことができる。

$$CL = \bar{X} \pm t \cdot S_x \quad (3.11)$$

で求めることができる。この場合, $df = 30$, $\alpha = 0.05$ の両側検定の t の値, $t = 2.042$ を利用する (統計法のテキストの巻末を参照), 例えば「時点1」の CL は

$$CL = 9.484 - 2.042 \times 0.334 = 8.802$$

$$CL = 9.484 + 2.042 \times 0.334 = 10.166$$

となる。この2つの値の範囲に, 母集団の平均値 μ が95%の確からしきで含まれることが示されている。

この値は, ひとつの標本 (例えば今回測定した1クラス31人) から母集団 (小3全体など) の様子を推定するひとつの目安になる。

この場合, このクラスの平均は9.484であったが, 母集団の平均は8.802から10.166のどこかに含まれると推定される。つまり, このくらいの幅を持たせて解釈しておいたほうがよいということでもある。

図17は, 平均値の推移グラフに, 95%信頼区間と, 多重比較の結果を合わせて表示したものである。

(5) 多重比較

「ペアごとの比較」から, 3時点間のどこに有意な差があるのかを検討することができる (図18)。

3時点以上の場合, 2時点ごとに対応のある t 検定を行うことは望ましくないと言われている (永田ら, 1997)。例えば, 3時点では全部で3つのペア比較を行いうるが, $\alpha = 0.05$ で対応のある t 検定を行うとき, H_0 が真のとき正しく H_0 を採択する確率, すなわ

被験者内効果の検定

測定変数名：MEASURE_1

ソース		タイプIII 平方和	自由度	平均平方	F 値	有意確率
factor 1	② 球面性の仮定	20.667	2	10.333	12.157	0.000
	Greenhouse-Geisser	20.667	1.995	10.360	12.157	0.000
	Huynh-Feldt	20.667	2.000	10.333	12.157	0.000
	下限	20.667	1.000	20.667	12.157	0.002
誤差 (factor 1)	③ 球面性の仮定	51.000	60	0.850		
	Greenhouse-Geisser	51.000	59.845	0.852		
	Huynh-Feldt	51.000	60.000	0.850		
	下限	51.000	30.000	1.700		

図15 被験者内効果の検定

推定値

測定変数名：MEASURE_1

factor 1	平均値	標準誤差	95%信頼区間	
			下限	上限
1	9.484	0.334	8.802	10.166
2	10.613	0.284	10.033	11.192
3	10.258	0.266	9.714	10.802

図16 95%信頼区間

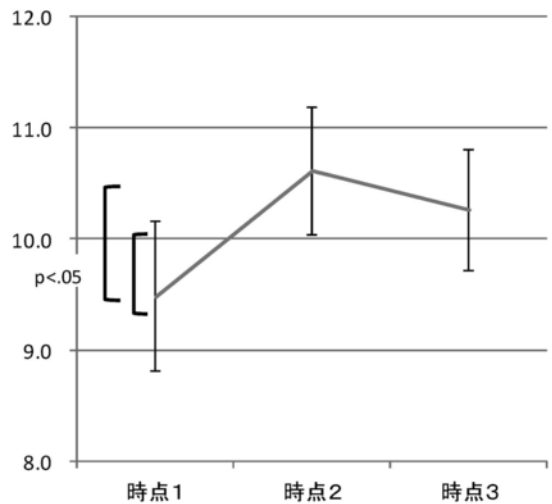


図17 平均値のグラフ

ち “間違わずに検定する確率” は

$$P=1-\alpha=1-0.05=0.95$$

であるから、これを 3 ペア分行うとすると、3 つとも
間違わずに検定する確率は

$$P'=0.95\times0.95\times0.95=0.857$$

となる。つまり、3 つのペア比較のうち少なくとも 1
つは間違える確率は

$$P=1-0.857=0.142$$

となり、当初設定した $\alpha=0.05$ を超えてしまうこと
になる。

そこで、1 回あたりの有意水準を調整し、全体とし
て当初設定した有意水準 α を超えないように調整
する。Bonferroni 法では、有意水準 α をペア比較の
数で割る。今回は 3 つの対比較を行うから

測定変数名：MEASURE_1

(I) factor1	(J) factor1	平均値の差 (I-J)	標準誤差	有意確率 ^b	95%平均差信頼区間 ^a	
					下限	上限
1	2	-1.129*	0.230	0.000	-1.711	-0.547
	3	-0.774*	0.240	0.009	-1.382	-0.166
2	1	1.129*	0.230	0.000	0.547	1.711
	3	0.355	0.233	0.416	-0.237	0.946
3	1	0.774*	0.240	0.009	0.166	1.382
	2	-0.355	0.233	0.416	-0.946	0.237

推定周辺平均に基づいた *、平均の差は 0.05 水準で有意です。

b. 多重比較の調整：Bonferroni。

図18 ペアごとの比較

$$\alpha = 0.05 / 3 = 0.0167$$

としてそれぞれのペア比較を行う。

図 17「平均値の差」*印がついているペアは、「有意差がある」ことを表している。今回は「時点 1」と「時点 2」, 「時点 1」と「時点 3」の間に有意差があることが分かる。

(6) 基礎的なデータを示すことの必要性

分散分析の結果は分散分析表によって表すのが原則であるが、変数が多い場合には、表 4 のように F および有意差の有無のみを要約して掲載することもある。この場合、平均値や SD 、データ数など基礎統計量や、検定の方法や df 、可能なら粗データ(匿名化された数値データ)についても、本文や付表など、報告書のどこかに含めておくことが望ましい。

近年、有意差だけでなく、効果の大きさ(効果量)

や区間推定の結果も求められることがある(APA, 2009; 大久保ら, 2012 など)。専門的な内容となるので本論では省くが、こうした統計量が求められなくとも、あとで読者が報告されている情報を元にして再計算できるよう、必要なデータは残しておくことが望ましい。

文献

APA 2009 Publication manual of the American Psychological Association (6th.ed): アメリカ心理学会(著), 前田・江藤・田中(訳) 2011 APA 論文作成マニュアル 第2版 医学書院。

新井啓子 2004 説明文の読みにおける『つなぎ言葉』に着目させた指導の効果, 聖徳大学大学院修士論文(In: 前原聡 2012 小学校国語科における「読むこと」に生かす文法指導の考察一文脈に沿って接続語を利用する「読み」を目指して一, 平成 23 年度群馬大学教職大学院課題研究報告書)

藤崎春代・武藤 隆 1985 幼児の協同遊びの構造—積木遊

表 4 小学生用 SEL-8S 自己評定尺度得点の比較

	時点 1		時点 2		時点 3		F
	平均値	SD	平均値	SD	平均値	SD	
自己への気づき	10.19	1.82	11.11	1.45	10.94	1.48	4.84*
他者への気づき	8.77	1.75	10.18	1.75	10.03	1.43	14.65**
自己のコントロール	9.39	1.76	9.60	1.78	9.71	1.72	0.62
対人関係	9.48	1.86	10.61	1.58	10.26	1.48	12.16**
責任ある意思決定	8.90	2.10	9.65	1.96	9.84	1.68	5.70**
生活上の問題防止のスキル	11.19	1.08	11.24	1.30	11.61	0.62	2.58+
人生の重要事態に対処する能力	8.87	2.03	9.58	1.82	9.29	1.94	2.15
積極的・貢献的な奉仕活動	9.42	1.78	10.26	1.91	10.48	1.84	7.18**

** : $p < 0.01$, * : $p < 0.05$, + : $p < 0.10$

びの場合一、教育心理学研究, 33(1), 33-42.
 群馬県教育委員会 2011 ぐんまの子どもの基礎・基本習得
 状況調査.
 服部 環・海保博之 1996 Q&A 心理データ解析. 福村出
 版.
 平石賢二 1990 自己肯定意識尺度. 堀洋道監修, 心理測定
 尺度集 I, サイエンス社, 16-22.
 犬塚美輪 2002 説明文における読解方略の構造. 教育心理
 学研究, 50(2), 152-162.
 小泉令三・山田洋平 2011 社会性と情動の学習 (SEL-8S)
 の進め方—小学校編. ミネルヴァ書房.
 国立教育政策研究所 2005 平成 15 年度小中学校教育課程
 実施状況調査.
 国立教育政策研究所 2006 特定の課題に関する調査 (算
 数・数学).
 国立教育政策研究所 2007 特定の課題に関する調査 (理

科).
 小谷野亘・長田久雄 1992 実証研究の手引き—調査と実験
 の進め方・まとめ方. ワールドプランニング. 第 6 章.
 永田 靖・吉田道弘 1997 統計的多重比較法の基礎. サイ
 エンティスト社. 第 2 章, 9-18.
 中村 修・宮崎 藍・渡邊美穂 2007 食育活動の現状と課
 題, 長崎大学総合環境研究, 10(1), 11-16.
 大久保街亜・岡田謙介 2012 伝えるための心理統計. 勁草
 書房.
 佐藤浩一 (編) 2013 学習の支援と教育評価—理論と実践
 の協同. 北大路書房.
 佐藤佐和子・児島秀貴 2007 内省を促し協働への意識を高
 める話合いの工夫—『話合いモデル』の作成と活用を通し
 て一. 群馬県総合教育センター研究報告書 平成 19 年度
 第 239 集. (In: 松原孝志 2011 学校における同僚性・
 協同性の形成に関する一考察—教師が語り合う校内研修

付表 1 「事前・事後の 2 時点の比較」のデータ

No.	事前合計	事前 1	事前 2	事前 3	事後合計	事後 1	事後 2	事後 3
1	27	3	4	3	45	3	4	4
2	36	4	4	4	47	4	4	4
3	40	4	4	4	42	3	4	3
4	37	3	4	2	41	4	4	2
5	42	3	4	2	51	4	4	4
6	37	3	4	1	45	3	3	4
7	41	3	4	3	39	3	4	2
8	38	4	4	3	46	3	4	4
9	25	2	2	2	41	3	4	4
10	37	1	2	3	44	4	3	4
11	45	3	3	4	42	4	2	4
12	28	2	3	2	30	3	4	3
13	50	4	4	4	50	4	4	4
14	42	3	3	3	31	1	1	1
15	46	4	4	4	49	4	3	4
16	49	4	4	4	49	4	4	4
17	48	4	4	4	52	4	4	4
18	46	4	4	4	52	4	4	4
19	38	3	4	3	46	3	4	3
20	30	3	4	3	36	3	3	4
21	42	4	4	3	46	3	4	4
22	39	3	4	1	52	4	4	4
23	40	4	4	3	52	4	4	4
24	35	2	4	1	47	4	4	4
25	39	4	4	3	49	4	4	3
26	33	3	4	2	49	4	4	4
27	27	3	3	2	38	3	2	3
28	44	2	4	4	43	4	4	3
29	26	2	3	1	45	4	4	4
30	39	3	4	2	39	3	2	2

* Excel に入力すると, セル範囲は A1 から I31 までになる。

- の推進を通して～,平成 22 年度群馬大学教職大学院課題研究報告書)
- 鋤柄増根 2002 研究法の理解とデータ分析における学生の誤解,教育心理学年報,41,104-113.
- 高橋志保 2013 児童の社会性を育む指導の工夫—コミュニケーションスキルの定着をめざしたカリキュラムの見直しを通して—平成 24 年度群馬大学教職大学院課題研究報告書.
- 多鹿秀継・中津梢男 2009 算数問題解決と転移を促す知識構成の研究,風間書房.
- 登張真稲 2003 青年期の共感性の発達:多次元的視点による検討,発達心理学研究,14(2),136-148.
- 山口陽弘 2013 教育評価における統計学的知識の重要性,In:佐藤浩一(編) 学習の支援と教育評価—理論と実践の協同,北大路書房,第 7 章,145-171.
- 山内光哉 2009 心理・教育のための統計法第 3 版,サイエンス社.
- 安永 悟 2006 実践・LTD 話し合い学習法,ナカニシヤ出版.

付表 2 「3 時点以上の比較」のデータ

No.	時点 1	時点 2	時点 3
1	6	6	8
2	6	11	9
3	11	11	9
4	8	12	11
5	10	12	12
6	11	10	11
7	9	10	9
8	10	12	9
9	8	10	8
10	12	12	12
11	10	10	12
12	7	8	10
13	9	12	11
14	12	12	12
15	7	7.5	8
16	12	12	12
17	10	11	11
18	11	12	12
19	12	12	12
20	12	12	11
21	9	10	8
22	7	9	9
23	11	12	12
24	12	12	11
25	8	9	8
26	10	12	10
27	10	11	12
28	8	9	10
29	9	9.5	10
30	8	10	10
31	9	11	9

* Excel に入力すると,セル範囲は A1 から D32 までになる。